Die Paradoxie der überraschenden Übung: Logische, epistemologische und pragmatische Aspekte

WOLFGANG LENZEN

1. Vor einem guten Vierteljahrhundert stellte D.J. O'Connor die folgende Paradoxie vor. Ein Offizier kündigt seiner Kompanie an, in der kommenden Woche werde eine *Überraschungsübung* stattfinden, die in Militärkreisen als Übung definiert ist, „von der die Teilnehmer nicht vor 6 Uhr am Morgen des Tages, an dem sie stattfindet, wissen können, dass sie [dann] stattfindet". Das Paradoxe hierbei ist, dass wegen der Bekanntmachung angeblich keine Überraschungsübung stattfinden kann, denn die Rekruten argumentieren wie folgt:

> Sie kann nicht am Samstag stattfinden, denn wenn sie an keinem der ersten fünf Tage jener Woche stattgefunden hat, muss sie am letzten stattfinden. Und die Tatsache, dass [wir] dies wissen können, verletzt die Bedingung, die [die Übung] definiert. Weil sie nicht am Samstag stattfinden kann, kann sie entsprechend auch nicht am Freitag stattfinden, denn wenn Samstag ausgeschlossen ist, ist Freitag der letzte mögliche Tag und kommt deshalb aus dem gleichen Grund nicht in Frage wie der Samstag. Durch entsprechende Argumente werden der Reihe nach Donnerstag, Mittwoch, usw. bis Sonntag ausgeschlossen, so dass die Übung überhaupt nicht stattfinden kann.[1]

Die Korrektheit dieses Arguments wurde in den ersten Diskussionen nie infrage gestellt: man diagnostizierte den vorliegenden Fall als Beispiel einer „pragmatically self-refuting announcement"[2], konzentrierte sich auf Formulierungsversuche einer Ankündigung, die sich nicht „selbst widerlegen" würde[3], und glaubte, die Paradoxie hiermit schon „erfolgreich ausgetrieben zu haben".[4]

Bald darauf bemerkte M. Scriven aber, dass ungeachtet der scharfsinnigen Rekrutenargumente dennoch eine Überraschungsübung stattfinden könne: wenn der Offizier etwa für den Mittwoch der fraglichen Woche eine Übung ansetzt, wird sie überraschend sein, denn offensichtlich konnte nicht gewusst werden, dass sie am Mittwoch stattfindet, oder – in Scrivens Worten – „the date of its occurrence cannot be forecast from the announcement".[5]

Diese Wendung war in der Tat „überraschend": Aus der Ankündigung einer Überraschungsübung scheint zunächst zu folgen, dass sie – gleich wann sie stattfindet – nicht überraschend ist, aber zugleich gilt anscheinend auch, dass sie doch (oder gar erst recht) überraschend ist, falls sie stattfindet.

Wenn die Korrektheit dieser beiden Argumente erwiesen wäre[6], läge also wirklich eine echte Paradoxie vor. Ob sie tatsächlich korrekt sind, soll im Folgenden untersucht werden. Dazu muss zunächst geklärt werden, was unter dem Begriff ‚überraschend' – und das heißt bei O'Connor: „nicht vor 6 Uhr …wissen können, dass..." – genau zu verstehen ist.

---

[1] Vgl. O'Connor [1948: S. 358].
[2] Vgl. O'Connor [1948], Cohen [1950] und Alexander [1950].
[3] Vgl. Cohen [1950] und Alexander [1950].
[4] Vgl. O'Connor [1951: S. 538].
[5] Scriven [1951: S. 403].
[6] Im Anhang weiter unten wird gezeigt, dass O'Connors Argument tatsächlich nicht schlüssig ist. Es ist verblüffend, wie unkritisch dieses Argument hier und dort übernommen wurde. A.K. Austin versucht z.B. in [1969] paradoxe Folgerungen dadurch zu vermeiden, dass er den Rekruten „praktische Anweisungen" gibt, um zu verhindern, „dass so eine Übung überraschend ist": Sie sollen irgendwann vor Montag den O'Connorschen Beweis bis zur Konklusion durchführen, dass die Übung Montag stattfinden muss, damit sie nicht überrascht werden; findet sie jedoch am Montag noch nicht statt, sollen sie sich vor Dienstag entsprechend klarmachen, dass sie Dienstag stattfinden muss, usw. Zur Naivität dieses Vorgehens ist wenig zu sagen. Nach Austins Anweisung müssten die Rekruten doch sehr überrascht sein, wenn die Übung etwa am Mittwoch stattfindet, da sie ja „bewiesen" hatten, dass sie am Montag (und außerdem am Dienstag) stattfinden muss!

Was heißt ‚wissen können' überhaupt, und aufgrund welcher Information kann dieses und jenes gewusst werden?

Die erste Frage wurde bislang hauptsächlich auf zwei Weisen beantwortet: ‚Wissen können, dass A' soll entweder bedeuten, dass A aus gewissen Prämissen logisch ableitbar ist, oder aber, dass aus gewissen Prämissen plus epistemologischen Prinzipien logisch ‚W(A)' folgt, wobei ‚W' ein geeignet einzuführender epistemischer Operator des Wissens ist. Die zweite Frage ist so zu beantworten, dass man die für die Herleitung zulässigen Prämissen jeweils angibt. Die so resultierenden logischen bzw. epistemologischen Varianten werden in den folgenden zwei Abschnitten diskutiert, während mein eigener, pragmatischer Ansatz im 4. Abschnitt dargestellt wird.

2. Die fragliche Paradoxie taucht in der Literatur in ganz unterschiedlichen Verkleidungen auf: von O'Connors militärischem Beispiel über den zivileren Fall einer überraschenden Klassenarbeit[7] bis hin zu Quines martialischer Fassung des „hangman". Wir wollen von diesen Versionen abstrahieren und ein neutrales Ereignis E betrachten, das – so lautet der erste Teil der jeweiligen Ankündigung – an genau einem der folgenden n Tage stattfindet. Die Sätze, dass E am j-ten Tag stattfindet ($j = 1,\ldots, n$), ließen sich in einer Prädikatenlogik 1. Stufe wiedergeben; wir wollen sie hier aber einfach durch die metasprachlichen Formeln $E(j)$ darstellen, so dass die gerade angegebene Teilankündigung symbolisch die Gestalt annimmt:

$$A: = \exists! jE(j)^{8},$$

wenn wir als Konvention den Quantor $\exists j$ (und später entsprechend $\forall j$) stets über $j = 1, ..., n$ laufen lassen.

Der Überraschungscharakter von E ist nach den obigen Ausführungen in der Form wiederzugeben: Wann immer E eintritt, ist aus den (bis dahin) zulässigen Prämissen P_j nicht logisch ableitbar, dass E dann stattfindet, symbolisch:

$$B := \forall j(E(j) \supset \neg(P_j \rightarrow E(j))).^{9}$$

Welche Prämissen sind nun (am j-ten Tag nach der Ankündigung) zulässig? Es ist klar, dass P_j auf jeden Fall den Satz A enthalten darf, wenn man sich an der informellen Fassung der Paradoxie orientiert. Weitere plausible Anwärter auf eine Mitgliedschaft in P_j sind jeweils die Sätze, dass, wenn $E(j)$, E bis zum j-1-ten Tag noch nicht stattgefunden hat, denn auch dies ist ein wesentlicher Bestandteil des Rekrutenarguments.[10] Übernimmt man

[7] Die am meisten diskutierte Variante – „Surprise Test", „Surprise Examination", „Unexpected Examination", usw. – hat zuerst Weiss [1952] angegeben; die Quinesche Fassung in [1953] erinnert ein wenig an die bekannte Paradoxie in Cervantes' *Don Quijote*, 2. Buch. Kap. 51.

[8] Wir verwenden die üblichen logischen Operatoren $\neg$, $\wedge$, $\vee$, $\supset$, $\equiv$, $\forall$ und $\exists$ als metasprachliche Zeichen für ‚nicht', ‚und', ‚oder', ‚wenn, dann', ‚genau dann, wenn', ‚für alle' bzw. ‚es gibt ein'; das Symbol ‚$\exists!x$' steht kurz für ‚Es gibt *genau* ein x, so dass'. Ferner dient $\rightarrow$ als Mitteilungszeichen für ‚folgt logisch' (in der zugrunde liegenden Prädikatenlogik). Unser Satz A entspricht dem Satz A in Bennett [1965].

[9] Bennett betrachtet in [1965] weitere „mögliche Interpretationen", die nicht mit dem Schema B übereinstimmen, sondern die Gestalt $\forall j(\neg(P_j \rightarrow E(j))$ haben. Diese Versionen werden jedoch kontradiktorisch, wenn P_j die „Minimalinformation" $A\wedge NT_{j-1}$ enthält.

[10] J. Schoenberg versucht die Paradoxie in [1966] dadurch zu lösen, dass sie die Prämissen $NT_{j\text{-}1}$ als unzulässig deklariert, da sie der Ankündigung angeblich widersprechen. Ein reductio-ad-absurdum Argument ist aber kein „logischer Fehlschluss", wie Schoenberg behauptet: $NT_{j\text{-}1}$ wird zum Beispiel nicht als „Prämisse", von deren Wahrheit man überzeugt wäre, sondern als hypothetische Annahme genommen, die gerade weil sie der Ankündigung widerspricht, als falsch erwiesen werden soll!

aus Bennett [1965] die Abkürzung NT_j für ‚E ist bis zum j-ten Tag noch nicht eingetreten', also

$$NT_j := \neg E(1)\wedge\neg E(2)\wedge\ldots\wedge\neg E(j)^{11}$$

so ergibt sich als eine erste logische Interpretation der Überraschungsankündigung

$$B_1 := \forall j(E(j) \supset \neg(A,\ NT_{j-1} \rightarrow E(j))).^{12}$$

Die Gesamtankündigung $A\wedge B_1$ ist dann nicht nur konsistent, sondern auch wahr, wenn E an irgendeinem der ersten n-1 Tage stattfindet. Der erste Schritt des O'Connorschen Arguments, mit dem der letzte Tag als möglicher Termin für E ausgeschlossen wird, wäre hier korrekt, denn trivialerweise gilt

$$A,\ NT_{n-1} \rightarrow E(n),$$

so dass aus B_1 für $j = n$ per modus tollens $\neg E(n)$ folgt. Der zweite Schritt ließe sich aber nicht mehr durchführen: Aus B_1 gewinnt man für $j = n$-1 zwar

$$E(n\text{-}1) \supset \neg(A,\ NT_{\text{n-2}} \rightarrow E(n\text{-}1)),$$

aber es gilt nur

$$A,\ NT_{\text{n-2}} \rightarrow E(n\text{-}1)\vee E(n),$$

und das aufgrund von B_1 schon ausgeschlossene $E(n)$ kann mit $A\wedge NT_{n-2}$ alleine nicht ausgeschlossen werden. Übertragen auf den informellen Gedankengang führt die Deutung von ‚überraschend' durch B_1 also zu folgender Situation: Obwohl in dem Fall, dass bis Donnerstag noch keine Übung stattgefunden hat, feststeht, dass die Übung nur am Freitag oder am Samstag abgehalten werden kann, und obwohl *aufgrund der Ankündigung* Samstag als möglicher Termin ausfällt, kann man *aufgrund der am Donnerstag zulässigen Information* nicht schließen, dass sie Freitag stattfindet. Sie wäre (aus der Donnerstag-Perspektive) im Sinne von B_1 überraschend, obwohl sicher ist, dass sie Freitag stattfindet.

Nach dieser ersten Analyse ist O'Connors Argument also fehlerhaft. Um es korrekt durchführen zu können, müsste man stärkere Prämissen zulassen. Da im informellen Gedankengang die Tatsache, dass E überraschend eintreten soll, ihrerseits dazu benutzt wird, weitere Überraschungstermine auszuschließen, bietet es sich an, B_1 selber als zulässige Prämisse in der Definition B_2 von ‚überraschend im 2. Sinn' aufzunehmen, so dass sich ergibt:

$$B_2 := \forall j(E(j) \supset\neg(\mathrm{A},\ NT_{j-1},\ B_1 \rightarrow E(j))).^{13}$$

Aus B_2 folgt dann $\neg E(n\text{-}1)\wedge\neg E(n)$, so dass die Ankündigung $A\wedge B_2$ im 2-Tage-Fall schon kontradiktorisch ist, ein Ergebnis, das man durch Hinzunahme von

$$B_3 := \forall j(E(j) \supset \neg(A,\ NT_{j-1},\ B_2 \rightarrow E(j)))$$
$$\ldots$$
$$B_n := \forall j(E(j) \supset \neg(A,\ NT_{j-1},\ B_{n-1} \rightarrow E(j)))$$

für $A\wedge B_{\text{n}}$ ebenfalls erzielt. Die gleiche Inkonsistenz ergibt sich auch, wenn man diese mehrstufige Definition in den einen „self-referring" Satz $B*$ komprimiert:

$$B*: = \forall j(E(j) \supset\neg(A,\ NT_{j-1},\ B* \rightarrow E(j))).^{14}$$

---

[11] NT_o sei eine Tautologie.

[12] B_1 entspricht Bennetts A_1 und dem (c') in Windt [1973]; vgl. auch die Regel $R2$ in Shaw [1958] und Lyon [1959].

[13] Vgl. Windt [1973: S. 68] sowie die Regel $R3$ in Shaw [1958]. B_2 ist ein meta-metasprachlicher Satz so dass $\rightarrow$ hier nicht mehr die logische Folgerungsbeziehung in der zugrunde liegenden Prädikatenlogik symbolisiert, sondern die logische Folgerung in der Metasprache.

Dieser Satz wurde bislang weitgehend als adäquateste Wiedergabe der Ankündigung angesehen. Vom logischen Standpunkt aus erscheint er aber suspekt, und man darf ihn sicher nicht als eine Definition oder Explikation von ‚überraschend' auffassen.[15] Im Hinblick auf die zu untersuchende Paradoxie können wir solche Skrupel jedoch außer Acht lassen, denn selbst mit B^* entsteht nichts Paradoxes: $A \wedge B^*$ ist logisch falsch, nicht mehr und nicht weniger. O'Connors Argument wird dadurch trivialerweise korrekt, Scrivens hingegen falsch, denn aus $A \wedge B^*$ folgt jeder Satz, also auch „das Datum des Eintretens von E"; wegen $\vdash \neg(A \wedge B^*)$ gilt $\vdash (A \supset \neg B^*)$, d. h. E wird eben *nicht* überraschend (im Sinne von B^*) sein, wenn es stattfindet.[16] Eine Paradoxie entsteht also auch hier nicht.

Echt paradoxe Ergebnisse erzielt man erst, wenn man nach einem Vorschlag von Cargile [1965] Ableitbarkeitsbehauptungen der Art ‚Aus A ist B logisch ableitbar' im Sinne von ‚$A \wedge (A \rightarrow B)$' interpretiert und das Schema B somit ersetzt durch

$$C := \forall j(E(j) \supset \neg(P_j \wedge (P_j \rightarrow E(j)))).^{17}$$

Für die kritische Wahl $P_j = \{A, NT_{j-1}, C\}$ ergibt sich nämlich

$$C^* = \forall j(E(j) \supset \neg(A \wedge NT_{j-1} \wedge C^* \wedge (A, NT_{j-1}, C^* \rightarrow E(j)))),^{18}$$

was für $n = 1$ zu

$$C^{**} = \neg(A \wedge C^{**})$$

führt. (Man beachte, dass NT_o tautologisch, A äquivalent mit $E(1)$ und $(A, C^{**} \rightarrow A)$ beweisbar ist!) C^{**} hat die bekannte Form des „conjunct liar", aus dem sich u. a. ergibt: Wenn C^{**} wahr ist, dann kann per „definitionem"[19] nicht sowohl A als auch C^{**} wahr sein, d. h. aus C^{**} „folgt" $\neg A$; deshalb „folgt" aus A gemäß Kontraposition $\neg C^{**}$, also auch $\neg(A \wedge C^{**})$, d. h. per „definitionem" auch C^{**}. Auf den ersten Blick könnten also sowohl das O'Connorsche als auch das Scrivensche Argument korrekt erscheinen: Falls E überraschend ist (C^{**}), wird E nicht stattfinden ($\neg A$), aber wenn E doch stattfindet (A), wird es überraschend sein (C^{**}).

Nun, aus A „folgt" aber nicht nur C^{**}, sondern zuallererst auch $\neg C^{**}$, so dass A als kontradiktorisch „bewiesen" ist. Die Scrivensche Konklusion „Würde E eintreten, so wäre es überraschend" wird bei der Deutung von ‚überraschend' durch C^{**} also ein irrealer Konditionalsatz mit „logisch falschem" Antezedenz, der natürlich nicht die Erfüllbarkeit der Ankündigung begründen kann und der zugleich zeigt, dass die Cargilesche Paradoxie mit dem ursprünglichen Rätsel, bei dem A selbstverständlich wahr sein kann, nichts zu tun hat.

---

[14] B^* entspricht z. B. Shaws $R2^*$ und Bennetts A_3. Vgl. auch Lyon [1959], Nerlich [1961], Fitch [1964] sowie Boschs mengentheoretische Version in [1972]. O'Carroll kommt in [1967] zu dem überraschenden Schluss, dass nicht nur B^*, sondern auch das B mit tautologischem P_j mit A logisch unverträglich ist. Dies resultiert freilich daher, dass er logische Implikation mit materialer Implikation verwechselt und B somit deutet als $\forall j(E(j) \supset \neg(P_j \supset E(j)))$. Nerlich hatte in [1961] schon mit der gleichen Formel operiert, aber (glücklicherweise) das rechte $\supset$ als $\rightarrow$ interpretiert.

[15] Dass B^* keine korrekte Definition oder Explikation sein kann, ist trivial; dass es aber auch sonst recht dubios ist, ergibt sich u. a. daraus, dass im Gegensatz zu B_n, wo $\rightarrow$ die logische Folgerung in der n-1-fachen Metasprache symbolisiert, bei B^* weder gesagt werden kann, zu welcher Sprachstufe es gehört, noch, auf welche Sprachstufe sich $\rightarrow$ bezieht.

[16] Natürlich folgt *aus der Ankündigung* auch, dass die Übung überraschend sein wird, wenn sie stattfindet, denn trivialerweise gilt $A \wedge B^* \rightarrow A \supset B^*$, aber das hat nichts mit einer Paradoxie zu tun. Nerlich operiert in [1961] mit den Folgerungen aus der anerkannt kontradiktorischen Ankündigung so lange herum, bis er glaubt, dass man gar nichts mehr wissen könne, und resignierend zugesteht: „This all makes one's head spin" (o.c., S. 512).

[17] Diesen Ansatz hat Windt mehr oder weniger wörtlich in [1973] übernommen.

[18] C^* entspricht dem Satz A_9 in Cargile [1965] und der Konjunktion von (A), (B) und (C) in Windt [1973: S. 67].

[19] Diese und entsprechende Anführungen im folgenden Text sollen daran erinnern, dass eben keine korrekten Definitionen vorliegen, und dass die „Folgerungen" und „Beweise" deshalb mit einiger Skepsis zu betrachten sind.

3. Wir wollen deshalb untersuchen, ob die epistemologischen (e.l.) Ansätze vielleicht zu einer adäquateren Analyse des O'Connorschen Rätsels führen. Dazu setzen wir eine e.1. Erweiterung der zugrunde gelegten Prädikatenlogik voraus, in der sich Sätze der Art ‚Vor dem Zeitpunkt i wird gewusst, dass A' darstellen lassen. Diese Sätze sollen hier wieder einfach durch metasprachliche Formeln W_iA wiedergegeben werden, und das angenommene e.l. System möge die folgenden Prinzipien enthalten:

$$W1\colon W_iA \supset A$$
$$W2\colon W_i(A \supset B) \supset (W_iA \supset W_iB)$$
$$W3\colon A \vdash W_iA.$$ [20]

Der Überraschungscharakter von E wäre dann, in enger Parallele zu den logischen Ansätzen, durch das Schema zu erfassen:

$$D := \forall j(E(j) \supset \neg(P_j \rightarrow W_jE(j))).$$

Mit der „Minimal"information $P_j = \{A, NT_{j\text{-}1}\}$ ist dann nicht einmal der erste Schritt des O'Connorschen Arguments durchführbar, denn aus

$$D_1 := \forall j(E(j) \supset \neg(A, NT_{j\text{-}1} \rightarrow W_jE(j)))$$

folgt für $j = n$

$$E(n) \supset \neg(A, NT_{n\text{-}1} \rightarrow W_nE(n));$$

aus $A \wedge NT_{n\text{-}1}$ kann man aber nur $E(n)$, und nicht $W_nE(n)$ erschließen. Um den Termin n per modus tollens ausschließen zu können, bräuchte man statt $A \wedge NT_{n\text{-}1}$ also die stärkere Prämisse $W_n(A \wedge NT_{n\text{-}1})$, d. h. $W_n(A) \wedge W_n(NT_{n\text{-}1})$. Der zweite Teil dieser Information soll durch das Prinzip des „Wissens aus der Erinnerung" gewährleistet sein, das ja auch implizit im Rekrutenargument benutzt wurde: Wenn die Übung am vorletzten Tag noch nicht stattgefunden hat, wüsste man es dann; formal:

$$NT_{n\text{-}1} \supset W_nNT_{n\text{-}1},$$ [21]

aber W_nA lässt sich nicht analog durch A begründen. Und genau hierin sieht Quine die Lösung der Paradoxie – „The puzzle is [...] to find the fallacy" –: Man dürfe zwar davon ausgehen, dass die (Teil-)Ankündigung A wahr ist, nicht aber davon, dass man wüsste, dass sie wahr ist, denn selbst aufgrund von Ankündigungen weiß man (im strengen Sinne) nie, was die Zukunft bringt.[22]

Nun, meiner Meinung nach entspricht es zwar durchaus dem normalen Sprachgebrauch, wenn man aufgrund einer *ordnungsgemäßen* Ankündigung eines Ereignisses E behauptet, man wüsste, dass E stattfindet.[23] Wenn ich etwa im Rundfunk die Ansage höre ‚Heute Abend um 20 Uhr überträgt der Bayerische Rundfunk ein Symphoniekonzert', dann fühle ich mich völlig berechtigt zu sagen, ich wüsste, dass ein Konzert *übertragen wird*, obwohl ich im strengen Sinne nur weiß, dass das Konzert *übertragen werden soll*, d. h. eben, dass zuverlässig angekündigt wurde, dass es übertragen wird.

Aber im vorliegenden Fall hat Quine mit seiner Kritik an der Annahme W_nA sicher Recht, denn der fragliche Ausspruch des Kommandanten ist alles andere als eine verlässliche, ord-

[20] Die in der epistemischen Logik übliche Relativierung des Operators W auf Personen(gruppen) braucht hier nicht explizit durchgeführt zu werden, da W sich im Folgenden stets auf ein und dieselbe Person(engruppe) bezieht. Das Ableitbarkeitszeichen $\vdash$ in der Regel $W3$, die der „rule of necessitation" in der Modallogik entspricht, muss von dem in diesem Abschnitt als ‚folgt epistemologisch' zu deutenden Zeichen $\rightarrow$ unterschieden werden, das sich wie folgt definieren lässt: $A \rightarrow B := \vdash (A \supset B)$.
[21] Vgl. etwa Kaplan/Montague [1960: S. 81] und Binkley [1968: S. 131].
[22] Vgl. Quine [1953: S. 65]; vgl: auch Weiss [1952: S. 269]: "[…] the last day cannot be dislocated from the other days at the time the announcement is made, except by a kind of theoretical anticipation of actual history." Zur Kritik der Quineschen Lösung vgl. Shaw [1958: S. 382], Medlin [1964: S. 66], Bosch [1972: § 2], Windt [1973: S. 67, Anm. 2] sowie Dietl [1973: S. 155].
[23] Vgl. dazu etwa Hintikka [1962: S. 90-91].

nungsgemäße Ankündigung. Das erweist sich u. a. daran, dass die Ankündigung – wie im Anhang weiter unten bewiesen wird – nicht „rational wissbar“ ist.

Davon abgesehen führt die fragwürdige Prämisse aber auch zu keiner neuen, interessanten Analyse der Paradoxie. Genau wie mit B_l lässt sich mit

$$D'_1 := \forall j(E(j) \supset \neg(W_j(A \wedge NT_{j-1}) \rightarrow W_j E(j)))$$

lediglich n als möglicher Termin für E ausschließen. Um weiterhin n-1 zu eliminieren, müsste ein entsprechendes D'_2 eingeführt werden, das nicht nur D'_1, sondern $W_{n-1}D'_1$ als zusätzliche Prämisse enthält,[24] und für das entsprechende D'_n bzw.

$$D^* = \forall j(E(j) \supset \neg(W_j(A \wedge NT_{j-1} \wedge D^*) \rightarrow W_j E(j)))$$

würde die Gesamtankündigung wieder inkonsistent. Diese Varianten brauchen also nicht noch einmal untersucht zu werden. Kaplan und Montague haben in [1960] schließlich einen weiteren Ansatz entwickelt, der auf den ersten Blick plausibler als die bisherigen Versuche erscheinen könnte. Sie interpretieren ‚E wird überraschend sein’ durch ‚Wann immer E eintritt, weiß man vorher *aufgrund der Ankündigung* nicht, dass E dann stattfindet’, und sie definieren diese Relation des Wissens eines Satzes A aufgrund eines Satzes B rein logisch durch $W(B \supset A)$.[25] Die Gesamtankündigung nimmt dann die symbolische Gestalt an:

$$F^* = A \wedge \forall j(E(j) \supset \neg W_j(F^* \wedge NT_{j-1} \supset E(j))).$$

Da sich F^* als kontradiktorisch herausstellte, schwächten die Autoren sie zu folgender Ankündigung ab:

$$G^* = \neg W_1 \neg G^* \supset A \wedge \forall j(E(j) \supset \neg W_j(G^* \wedge NT_{j-1} \supset E(j))),$$ [26]

die im 1-Tage-Fall zu

$$H^* = \neg W_1 \neg H^* \supset A \wedge \neg W_1(H^* \supset A)$$

und damit zu folgender Paradoxie führt: Wegen $W1$ „folgt“ aus H^* $\neg W \neg H^*$ (den Index bei W lassen wir der Einfachheit halber weg), und damit insbesondere auch A; wegen $W3$ gilt dann $W(H^* \supset A)$, so dass aus H^* auch $\neg H^*$ folgt und somit „beweisbar“ falsch ist; dann ist gemäß $W3$ aber auch $W(\neg H^{**})$, und deshalb per „definitionem“ H^* selber „beweisbar“, d. h. Kaplan/Montagues Ankündigung ist sowohl „beweisbar“ falsch als auch „beweisbar“ wahr!

Was bedeutet dies aber für das O’Connorsche Rätsel? H^* lautet umgangssprachlich etwa: ‚Wenn nicht gewusst wird, dass diese Ankündigung falsch ist, findet E morgen statt, ohne dass aufgrund dieser Ankündigung gewusst wird, dass E dann stattfindet’, also: ‚Wenn nicht gewusst wird, dass diese Ankündigung falsch ist, findet E morgen überraschend statt.’

Das O’Connorsche Argument ist also gültig, da „bewiesen“ wurde, dass E nicht überraschend eintritt ($W(H^* \supset A)$), und Scriven hat hier insofern Recht, als E „notwendigerweise“ überraschend stattfindet, da wegen des „bewiesenen“ $H^* \wedge \neg W \neg H^*$ auch $A \wedge \neg W(H^* \supset A)$ „beweisbar wahr“ ist. Trotzdem liegt hier keine adäquate Analyse des O’Connorschen Rätsels vor, denn zum einen enthält die informelle Ankündigung nicht die fragliche ‚wenn ... nicht’-Klausel, durch die sich allein die paradoxen Folgerungen ergeben, und zum anderen gilt im Widerspruch zur intuitiven Situation, dass das angekündigte Ereignis E aus „logischen“ Gründen stattfinden müsste.

---

[24] McLelland will in [1971] zeigen, dass nur der Termin *n* ausgeschlossen werden darf: Man dürfe zwar davon ausgehen, dass die Übung überraschend ist, nicht aber davon, dass man dies weiß, d. h. er hält – in unserer Terminologie – D_l, nicht aber $W_{n-2}D_1$ für eine zulässige Prämisse. Obwohl dies nicht ganz falsch ist – vgl. den Anhang weiter unten – führt seine Analyse zu dem merkwürdigen Ergebnis, dass *a* (die Person, die bei ihm das fragliche Argument durchführt), „does not know himself“ (o.c., S. 83).

[25] Dass dieser logische Ansatz prinzipiell unbefriedigend ist, wird weiter unten ausführlich begründet.

[26] Vgl. insbesondere die 2-Tage und 0-Tage-Versionen D_4 und D_5 in Kaplan/Montague [1960].

4. Überhaupt scheinen die bislang betrachteten Lösungsversuche aus mehreren Gründen unbefriedigend. Zum ersten hängt die Frage, ob ein angekündigtes Ereignis E für eine Person a überraschend ist, nicht allein vom Inhalt bzw. von der Bedeutung der Ankündigung ab: Wenn mir z. B. eine Wahrsagerin prophezeien würde, dass ich am nächsten Tag ein Säckchen Goldstücke unter meinem Kopfkissen finden werde, dann wäre ich – falls die Prophezeiung sich erfüllt – sehr überrascht, da ich es trotz der Vorhersage für praktisch ausgeschlossen hielte, dass das Gold unter dem Kopfkissen auftauchen sollte. Mein (fiktiver) abergläubischer Freund Hans, der mit der Wahrsagerin schon gute Erfahrungen gemacht haben möge, könnte dagegen überzeugt sein, dass sie auch diesmal die Wahrheit sagt, und wäre deshalb über das mögliche Gold keineswegs überrascht. Ob ein prophezeites oder sonst wie angekündigtes Ereignis überraschend ist, lässt sich also nur unter Bezug auf die pragmatischen Umstände der Ankündigung adäquat erfassen.

Zum zweiten erweist sich der bislang beschrittene Weg, die fraglichen Argumente auf ihre Gültigkeit hin zu untersuchen, anscheinend prinzipiell als Sackgasse. Wenn man O'Connors Konklusion α, es könne keine Überraschungsübung stattfinden, durch den Nachweis zu verifizieren versucht, dass α aus der entsprechend definierten Ankündigung β *logisch folgt*, so muss β zwangsläufig kontradiktorisch sein, da sie ja gerade besagt, dass eine Überraschungsübung stattfinden wird. Eine intuitiv adäquate Analyse des O'Connor/Scrivenschen Rätsels sollte deshalb überprüfen, ob α durch *den Akt der Ankündigung pragmatisch bewirkt* wird, und nicht, ob die Ankündigungsaussage β *logisch* α *impliziert*.[27]

Dazu muss zuerst klargestellt werden, was denn ‚überraschend' im umgangssprachlichen Sinn genau bedeutet. Es ist sicherlich inadäquat, ein Ereignis E für eine Person a überraschend zu nennen, wenn a lediglich *nicht weiß*, dass E eintreten wird: Ich weiß z. B. nicht, ob der nächste Wurf mit einer vorliegenden Münze „Kopf" oder „Zahl" ergeben wird, aber deshalb gilt keineswegs, dass ich überrascht sein werde, wenn das eine oder das andere Ergebnis eintritt; und dieses Beispiel zeigt auch, dass selbst die schärfere Bedingung: E tritt für a überraschend ein, wenn *a nicht glaubt*, dass E, immer noch zu schwach ist, denn bei dem zitierten Münzwurf glaube ich weder, dass „Kopf", noch, dass „Zahl" auftreten wird. Dem umgangssprachlichen Sinn kommt man wohl am nächsten, wenn man E genau dann für a überraschend nennt, wenn die subjektive Wahrscheinlichkeit von E für a so klein ist, dass a glaubt, dass E nicht eintreten wird.[28] Diese

---

[27] Auf diese pragmatische Dimension von Ankündigungen hat einzig Scriven explizit hingewiesen, der in [1952] zwischen einer Ankündigung als Bekanntmachung (*ordainment*) und als Aussage (*statement*) unterscheidet. Seine Lösung der Paradoxie bleibt jedoch unbefriedigend, weil er den Überraschungscharakter von E nicht durch die pragmatische Funktion der Bekanntmachung definiert, sondern durch die logischen Konsequenzen der Ankündigung.

[28] Wenn G_iA analog zu W_iA bedeutet ‚vor dem Zeitpunkt i wird geglaubt, dass A', so wäre ‚überraschend' im einen Fall durch $\forall j(E(j) \supset \neg G_jE(j))$, im anderen durch $\forall j(E(j) \supset G_j\neg E(j))$ wiederzugeben. (Vgl. dazu Harrisons Definitionen von ‚unforeseen', ‚unexpected' und ‚surprised' in [1969: S. 75]). Binkley stützt sich in [1968] auf die erste der beiden obigen Formeln und versucht dann die Inkonsistenz der entsprechenden 1-Tage-Ankündigung: $E(1) \wedge \neg G_1E(1)$ mit Hilfe eines der folgenden Prinzipien zu begründen: (vgl. o.c., S. 131)

$A5a$ $G_0A \supset G_1A$

$A5b$ $G_0A \supset G_0G_1A$

$A5c$ $G_0A \supset \neg G_0\neg G_1A$.

$A5a$ lässt sich als vernünftiges Prinzip für einen rationalen *Wissens*begriff ansehen (vgl. das Prinzip W_K im Anhang), aber sicher nicht für einen Begriff rationalen *Glaubens*. $A5b$ und $A5c$ sind, wie in Lenzen [1975] gezeigt werden soll, ebenfalls nicht unproblematisch. Dagegen gilt sicher das noch schwächere Prinzip $G_0A \supset \neg G_0G_1\neg A$, aus dem sich zusammen mit dem Analogon zur „KK-These" (vgl. etwa Hilpinen [1970]), $G_iA \supset G_iG_iA$, unmittelbar ergibt, dass die 1-Tage-Ankündigung in unserer Interpretation: $E(1) \wedge G_1\neg E(1)$ nicht „rational glaubbar" ist, d. h. dass $G_0(E(1) \wedge G_1\neg E(1))$ doxastologisch kontradiktorisch ist.

Überlegung führt im 1-Tage-Fall des O'Connorschen Rätsels zu folgender Lösung. Wenn wir davon ausgehen, dass Ankündigungen des Kommandanten normalerweise verlässlich sind, lässt sich zunächst feststellen, dass die Bekanntmachung (γ) ‚Morgen findet eine Übung statt' pragmatisch bewirkt, dass es keine Überraschungsübung wird, denn aufgrund der Bekanntmachung werden die Rekruten mit der Übung rechnen. Was ist aber, wenn verkündet wird: Morgen findet eine Überraschungsübung statt? (δ)

Zunächst einmal ist festzuhalten, dass δ keineswegs kontradiktorisch ist. δ wird genau dann wahr, wenn am nächsten Tag eine Übung stattfindet und wenn die Rekruten dann glauben, dass sie nicht stattfindet. δ ist aber sicher auch keine *verlässliche* Bekanntmachung, denn es ist doxasto-logisch unmöglich, damit zu rechnen, dass am nächsten Tag eine Übung stattfindet, von der man dann glaubt, dass sie nicht stattfindet: δ ist also nicht rational glaubbar und in diesem Sinne unzuverlässig.[29] Daraus alleine folgt allerdings noch nicht, ob die Übung – wenn sie stattfindet – überraschend sein wird oder nicht. Man kann nur sagen: wenn die Rekruten wegen der Unzuverlässigkeit der Ankündigung glauben, dass keine Übung stattfinden wird, dann wäre sie doch überraschend, d. h. in diesem Fall wäre Scrivens Argument korrekt.

Wenn die Rekruten dagegen trotz der erwiesenen Unzuverlässigkeit mit einer Übung rechnen (bzw. zumindest nicht glauben, dass sie nicht stattfindet), dann kann sie nicht überraschend sein, d. h. hier hätte O'Connor Recht: Obwohl aus δ *logisch folgt*, dass eine Überraschungsübung stattfindet, *bewirkt* die Bekanntmachung von δ, dass keine Überraschungsübung stattfindet.

Dies ist keine echte Paradoxie. Das einzig Paradoxe könnte man darin erblicken, dass im ersten Fall einer unzuverlässigen Ankündigung so sehr misstraut wird, dass sie dadurch (partiell) verifiziert wird, während im zweiten Fall das Vertrauen in die Ankündigung bewirkt, dass sie falsch wird, d. h. dass δ – welche Haltung die Rekruten auch immer einnehmen – entweder teilweise „pragmatisch verifiziert" oder aber „pragmatisch falsifiziert" wird.

Betrachten wir abschließend noch den 2-Tage-Fall, der schon alle wesentlichen Merkmale des Ausgangsrätsels enthält. Wenn der Kommandant bekannt gibt (ε) ‚Morgen oder übermorgen findet eine Übung statt', dann werden die Rekruten – die allgemeine Zuverlässigkeit ihres Vorgesetzten vorausgesetzt – mit der Übung rechnen und deshalb weder überrascht sein, wenn sie am ersten, noch, wenn sie erst am zweiten Tag stattfindet: Da die Ausgangswahrscheinlichkeit für beide Termine gleich groß ist, besteht offensichtlich kein Grund, den ersten Tag von vornherein auszuschließen; und wenn die Übung am ersten Tag noch nicht stattgefunden hat, besteht (nach allem, was zu erwarten ist) ebenfalls kein Grund, zu glauben, dass sie nun gar nicht mehr stattfindet.

Wenn der Kommandant aber ankündigt: ‚Morgen oder übermorgen findet eine Überraschungsübung statt' (η), könnten die Rekruten à la O'Connor überlegen:

> Würde die Übung erst am zweiten Tag stattfinden, so wären wir – falls wir dann glauben, dass er sein Versprechen hält – nicht überrascht; deshalb müsste die Übung eigentlich schon am ersten Tag stattfinden und wäre aus diesem Grunde ebenfalls nicht überraschend. Deshalb kann eigentlich überhaupt keine Überraschungsübung stattfinden.

Diese Überlegung ist aber irreführend. Die Rekruten sollten doch wissen, dass die Frage, ob die Übung überraschend wird oder nicht, einzig und allein von ihren Erwartungen abhängt. Wenn sie sich durch die merkwürdige Ankündigung oder durch das gerade angegebene Argument dazu verleiten lassen würden, nun überhaupt nicht mehr mit

---

Sowohl Binkley als auch Harrison beschäftigen sich in den genannten Arbeiten ausführlich mit dieser Frage der Konsistenz bzw. rationalen Glaubwürdigkeit der einschlägigen Ankündigungen, aber leider vergessen sie beide die Resultate ihrer Überlegungen auf das O'Connorsche Rätsel anzuwenden.

[29] Vgl. dazu Slaters kurze Bemerkungen in [1974].

einer Übung zu rechnen, so würden sie, falls sie doch stattfindet, überrascht. Wenn sie sich dagegen so verhalten, als hätte der Kommandant nie etwas von „überraschend“ gesagt, dann kann tatsächlich keine Überraschungsübung stattfinden.

Anhang:

Gegen die gerade dargestellte Lösung des O'Connorschen Rätsels könnte eingewendet werden, dass ‚überraschend’ bei O’Connor eben nicht im umgangssprachlichen Sinn von ‚glauben, dass nicht’ sondern im schwächeren Sinn von ‚nicht wissen, dass’ verstanden werden soll. Deshalb soll hier abschließend gezeigt werden, wo bei dieser Interpretation von ‚überraschend’ genau der Fehlschluss des informellen Arguments auftritt. Definieren wir ‚*E* wird überraschend sein’ durch ‚Wann immer *E* eintritt, weiß man vorher nicht, dass *E* dann eintritt’, also:

$$D^+ := \forall j(E(j) \supset \neg W_j E(j)),$$

so lautet die Ankündigung im 2-Tage-Fall:

$$E(1)\wedge\neg W_1 E(1) \vee E(2)\wedge\neg W_2 E(2).$$ [30]

Der erste Schritt des O'Connorschen Arguments wäre dann folgendermaßen zu symbolisieren:

1. $E(2)$ (Annahme)
2. $\neg E(1)$
3. $W_2 \neg E(1)$
4. $W_2 E(2)$
5. $\neg W_2 E(2)$

Also: $\neg E(2)$.

Für den Schritt von 1. auf 2. benötigt man, dass entweder $E(1)$ oder $E(2)$, d. h. dass A gilt; für den Schritt auf 3. braucht man das oben erwähnte Prinzip des „Wissens aus der Erinnerung“, kurz W_E; der Schritt auf 4. ist nur gültig, wenn $W_2(E(1)\vee E(2))$, also $W_2 A$ gilt, das man aus der von Quine kritisierten Prämisse $W_0 A$ mit Hilfe eines Prinzips der „Wissenskonservierung“

$$W_K : W_i B \supset W_j B \text{ für alle } i < j$$

gewinnt; und für den Schluss auf 5. braucht man schließlich D^+, so dass insgesamt gilt:

$$A\wedge D^+,\ W_0 A \rightarrow_W \neg E(2),$$

und somit auch

$$A\wedge D^+,\ W_0 A \rightarrow_W E(1),$$

wobei ‚$\rightarrow_W$’ bedeutet: ‚Folgt mit Hilfe der epistemologischen Prinzipien $W1$-3, W_E und W_K’.

Der nächste Schritt, mit dem $E(1)$ ausgeschlossen werden soll, ist jedoch nicht mehr durchführbar, denn $W_1 E(1)$ kann man epistemologisch nicht aus $W_0(A)$ und D^+ erschließen, sondern nur aus $W_0(A\wedge D^+)$!

Die O’Connorsche Ankündigung $A\wedge D^+$ ist also durchaus konsistent, aber nicht rational wissbar; das informelle Argument lässt sich nur unter der Voraussetzung korrekt durchführen, man wüsste, dass die Ankündigung wahr ist: Diese Voraussetzung ist aber beweisbar falsch.

[30] Das Zeichen ‚$\vee$’ symbolisiere hier das ausschließende ‚oder’.

Literatur

ALEXANDER, P.: „Pragmatic Paradoxes“. *Mind* 59 (1950), S. 536-538.
AUSTIN, A.K.: “On the Unexpected Examination”. *Mind* 78 (1969), S. 137.
BENNETT, J.: *Sammelrezension von* SHAW [1958], Lyon [1959], Nerlich [1961], Fitch [1964] und Medlin [1964. *The Journal of Symbolic Logic* 30 (1965), S. 101-102.
BINKLEY, R.: “The Surprise Examination in Modal Logic”. *The Journal of Philosophy* 65 (1968), S. 127-136.
BOSCH, J.: “The Examination Paradox”. *Logique et Analyse* 15 (1972), S. 505-525.
CARGILE, J.: *Sammelrezension von* Kaplan/Montague [1960] u. a. *The Journal of Symbolic Logic* 30 (1965), S. 102-103.
COHEN, L. J.: “Mr. O’Connor’s ‘Pragmatic Paradoxes’”. *Mind* 59 (1950), S. 85-87.
DIETL, P. J.: “The Surprise Examination”. *Educational Theory* 23 (1973), S. 153-158.
FITCH, F.B.: “A Goedelized Formulation of the Prediction Paradox”. *American Philosophical Quarterly* 1 (1964), S. 161-164.
HARRISON, C.: “The Unanticipated Examination in View of Kripke’s Semantics for Modal Logic”. In J. W. Davis et al. (eds.) *Philosophical Logic*. Dordrecht, Holland, 1969.
HILPINEN, R.: “Knowing that One Knows and the Classical Definition of Knowledge”. *Synthese* 21 (1970), S. 109-132.
HINTIKKA, J.: *Knowledge and Belief*. Ithaca, N.Y., 1962.
KAPLAN, D. UND R. MONTAGUE.: “A Paradox Regained”. *Notre Dame Journal of Formal Logic* 1 (1960), S. 79-90.
LENZEN, W.: “Probabilistic Interpretations of Epistemic Concepts”. *International Congress of Logic, Methodology, and Philosophy of Science*, 1975, Contributed Papers, S. IV 18-20.
LYON, A.: “The Prediction Paradox”. *Mind* 68 (1959), S. 510-517.
MCLELLAND, J.: “Epistemic Logic and the Paradox of the Surprise”. *International Logic Review* 3 (1971), S. 69-85.
MEDLIN, B.: “The Unexpected Examination”. *American Philosophical Quarterly* 1 (1964), S. 66-72.
NERLICH, G. C.: “Unexpected Examinations and Unprovable Statements”. *Mind* 70 (1961), S. 503-513.
O’CARROLL, M. J.: “Improper Self-Reference in Classical Logic and the Prediction Paradox”. *Logique et Analyse* 10 (1967), S. 167-172.
O’CONNOR, D. J.: “Pragmatic Paradoxes”. *Mind* 57 (1948), S. 358-359.
O’CONNOR, D.J.: “Pragmatic Paradoxes and Fugitive Propositions”. *Mind* 60 (1951), S. 536-538.
QUINE, W. V.: “On a So-Called Paradox”. *Mind* 62 (1953), S. 65-67.
SCHOENBERG, J.: “A Note on the Logical Fallacy in the Paradox of the Unexpected Examination”. *Mind* 75 (1966), S. 125-127.
SCRIVEN, M.: “Paradoxical Announcements”. *Mind* 60 (1951), S. 403-407
SHAW, R.: “The Paradox of the Unexpected Examination”. *Mind* 67 (1958), S. 382-384.
SLATER, B. H.: “The Examiner Examined”. *Analysis* 35 (1974), S. 49-50.
WEISS, P.: “The Prediction Paradox”. *Mind* 61 (1952), S. 265-269.
WINDT, P.: “The Liar in the Prediction Paradox”. *American Philosophical Quarterley* 10 (1973), S. 65-68.